

# Vývoj daňových úniků v ČR: Analýza pomocí markovských řetězců

Jan HANOUSEK\* – Filip PALDA\*\*

## 1. Motivace

Většina výzkumů, které se zabývají daňovými úniky, se zaměřuje na velmi populární měření velikosti šedé ekonomiky jako počtu neplatičů daní (pro přehled viz (Schneider, 2000)). Toto měření je však samo o sobě velmi problematické, zejména proto, že jeho postupy jsou často založeny na silných předpokladech a mezi jednotlivými aplikacemi převládají hlavně metody založené na makroekonomických agregátech (spotřeba elektrické energie, množství oběživa apod.).

Výzkumy založené na výběrových šetřeních se také setkávají s mnoha problémy při měření aktuální hladiny šedé ekonomiky v konkrétní zemi, protože odpovědi v dotaznících podchycují dolní hranici nepřiznaných příjmů a lze očekávat, že míra šedé ekonomiky bude ve skutečnosti vyšší. Pokud však máme k dispozici výsledky dotazování za několik let, můžeme dobře pracovat se změnami v úrovních šedé ekonomiky. Rozdíly či porovnání jednotlivých let ve výběrových šetřeních by měly odrážet příslušné změny v dané zemi. Podmínkou samozřejmě je, aby daný soubor byl reprezentativním výběrem a aby příslušná výběrová šetření byla provedena konzistentně podle stejné metodiky.

Je třeba si uvědomit, že postoj k šedé ekonomice se může měnit nejen v závislosti na změnách obecných ekonomických ukazatelů (ekonomický růst, daňové zatížení a související pokuty a efektivita daňových úřadů), ale také v souvislosti se změnami na úrovni jednotlivce/rodiny (ekonomická situace rodiny, vnímání využití daní, riziko dopadení, apod.).<sup>1</sup> Velikost šedé ekonomiky – a tím i úroveň vládních příjmů – se mění podle toho, do jaké míry se mění zapojení jednotlivců v ní. Kdybychom znali pravděpodobnost skutečnosti, že osoba s danými charakteristikami zahájí nebo ukončí svoji čin-

\* CERGE-EI, společné pracoviště Univerzity Karlovy a Akademie věd ČR (Jan.Hanousek@cerge-ei.cz) – korespondující autor

\*\* Ecole Nationale d'Administration Publique, Montreal, Canada (Filip\_Palda@enap.quebec.ca)

Autoři děkují Randallu Filerovi a Štěpánu Jurajdovi za připomínky k dřívější verzi tohoto článku. Tento výzkum vznikl za podpory Grantové agentury České republiky (402/04/0166), dřívější data byla získána za přispění grantů Phare ACE programu Evropské unie (P97-8119-R) a Nadace Volkswagen (II/75831). Autoři děkují za připomínky účastníkům řady seminářů uskutečněných na CERGE-EI, účastníkům Česko-francouzského ekonomického fóra, Paříž, 2004 a Atlantické ekonomické společnosti v Chicagu 2004.

<sup>1</sup> (Hanousek – Palda, 2004) obsahuje analýzu jednotlivých faktorů ovlivňujících účast jednotlivce v šedé ekonomice, která bere v potaz vliv kvality vládních služeb.

nost v šedé ekonomice, mohli bychom velmi dobře odhadnout výhledy do budoucna a vysledovat velikost a složení skupiny lidí, kteří jsou v šedé ekonomice aktivní. Ke zjištění této pravděpodobnosti nepotřebujeme primárně znát jejich důvody. Můžeme se spolehnout na časové řady participace v šedé ekonomice a použít je k odhadu toho, jak se bude vyvíjet chování jednotlivců v následujících letech.

K modelování a analýze přechodů mezi šedou a oficiální ekonomikou používáme nepanelová výběrová šetření provedená v České republice v letech 2000, 2002 a 2004. Ptali jsme se respondentů, zda se účastní (a účastnili) aktivit v šedé ekonomice – nikdy, někdy nebo často –, a sjednotili jsme tyto subjektivní odpovědi jen do dvou kategorií: aktivity v šedé ekonomice versus aktivity v oficiální ekonomice. Využili jsme při tom retrospektivního dotazování. Respondentů z roku 2000 jsme se ptali na jejich daňové úniky v letech 1995, 1999 a 2000, v roce 2002 jsme se ptali na roky 1997, 2000 a 2002 a respondentům v roce 2004 jsme položili stejnou otázku za roky 1999, 2002 a 2004. Na základě odhadu změn v dynamice šedé ekonomiky v minulosti můžeme usuzovat na její vývoj v následujících letech. Naše výsledky naznačují, že v současné době se daňové úniky v České republice stabilizovaly a během následujících pěti let lze očekávat mírný pokles neplacení daní.

Naše metodologie je nová v několika ohledech.

Za prvé, pomocí statistických testů ukazujeme, že naše výběrová šetření jsou navzájem konzistentní a že retrospektivní dotazování lze využívat ke studiu dynamiky vývoje šedé ekonomiky v České republice. Konzistence je zde chápána tak, že odpovědi respondentů na otázky v dřívějších průzkumech nesmějí být statisticky odlišné od odpovědí na stejné otázky v průzkumech pozdějších. Například v roce 2004 jsme se ptali respondentů na jejich daňové úniky v roce 2002, což je retrospektivní otázka. Retrospektivní odpovědi respondentů z průzkumu v roce 2004 nebyly statisticky rozlišitelné od odpovědí, které na tutéž otázku (tj. daňové úniky v roce 2002) dali respondenti v roce 2002 (současná otázka). Tento výsledek bereme jako potvrzení dvou potřebných předpokladů: a) výběrová šetření byla provedena stejnou metodologií a b) respondenti si závažnou skutečnost (daňový únik) dobře pamatují. Lze proto využívat retrospektivní otázky a navíc spojit náhodné vzorky z různých let a analyzovat je později jako jeden celek. Možnost spojit náhodné výběry je důležitá, protože chceme analyzovat činitele ovlivňující šedou ekonomiku pomocí regresních metod a spojování nám poskytne dostatečný počet pozorování k určení statisticky významných vztahů mezi šedou ekonomikou a proměnnými, které ji ovlivňují.

Druhým významným novým aspektem naší studie je fakt, že testujeme stabilitu markovských pravděpodobností. Tyto pravděpodobnosti by se *ceteris paribus* neměly měnit. V průběhu transformace je však v ekonomice jen málo věcí, které zůstanou *ceteris paribus*. Demografické charakteristiky, instituce i vládní politiky vztahující se k šedé ekonomice se velmi rychle mění. Změny nacházíme také v našich markovských pravděpodobnostech a připisujeme je uvedeným změnám parametrů, jež ovlivňují šedou ekonomiku.

V našem článku nejprve rozebereme použitá data tvořící základ analýzy a ukážeme, jak z nich lze odhadnout markovské pravděpodobnosti přechodu. Data pak použijeme k odhadu strukturálního a redukováného modelu těchto

pravděpodobností. Ukážeme, jak lze postup předpovědi spojit s různými scénáři vývoje faktorů, jež ovlivňují dynamiku šedé ekonomiky. Na základě předpokladů o vývoji zmíněných faktorů si pak budeme schopni udělat lepší představu o vývoji šedé ekonomiky v budoucnosti, případně představu o tom, do jaké míry může určitá hospodářská politika úroveň šedé ekonomiky ovlivnit.

## 2. Data za Českou republiku

Naše data pocházejí ze tří průzkumů provedených v České republice v letech 2000, 2002 a 2004. Souhrnné statistiky všech proměnných, které jsou v těchto průzkumech obsaženy, stejně tak jako použité otázky jsou k dispozici v článku (Hanousek – Palda, 2002).<sup>2</sup> Naše průzkumy jsou podobné průzkumům uvedeným ve (Fortin et al., 2000); technikou těchto autorů bylo dotazování (v našem případě osobní) zaměřené na získání informací ovlivňujících účast lidí v šedé ekonomice. Zmiňovaný výzkum se od našeho lišil tím, že neobsahoval otázky, které by umožnily odhadovat dynamiku vývoje šedé ekonomiky. Zajímal se pouze o souvislost mezi nákupem nezdaněných výrobků a služeb a daňovými úniky obecně.

Detailní analýzu našeho souboru dat uvádíme jinde (Hanousek – Palda, 2002); hlavními faktory, jež ovlivňují šedou ekonomiku, však je, že (a) míra participace závisí na věku (do středního věku podíl lidí v šedé ekonomice stoupá, poté klesá), (b) že šedá ekonomika je do velké míry doménou mužů a (c) že velmi často souvisí s prací na částečný úvazek a s nezaměstnaností. *Tabulky A1a–c v Příloze* uvádějí celkový přehled popisných statistik. Tyto tabulky obsahují jak redukované proměnné (věk, pohlaví), tak strukturní proměnné (vnímaná kvalita služeb poskytovaných vládou, subjektivní pravděpodobnost přistížení při účasti v šedé ekonomice). V závěrečné části tohoto článku budeme podrobněji rozebírat faktory, jež ovlivňují aktivitu jedince v šedé ekonomice, za použití nelineárního redukovaného regresního modelu.

### 2.1 Trendy ve vývoji šedé ekonomiky

Je nutné si uvědomit, že přiznání k daňovému úniku je de facto přiznání k nezákonnému jednání, za které může být jednatel potrestán. Pochopitelně musíme očekávat, že odhady úrovně šedé ekonomiky spočítané na základě našich dotazníků budou pouze dolním odhadem skutečnosti. Domníváme se, že mnohem zajímavější než odhad úrovně šedé ekonomiky je studie její dynamiky v čase. Proto hlavním předmětem našeho zájmu v tomto článku bude analýza aktivit jednotlivců v šedé ekonomice v čase a analýza faktorů, které tyto změny provázejí. K tomuto účelu jsme zařadili do dotazníku sérii (i retrospektivních) otázek, ve kterých jsme se respondentů ptali, jak často byli a jsou aktivní v šedé ekonomice. Dali jsme jim možnost odpovědět „velmi často“, „občas“ nebo „nikdy“. Zjistili jsme, že příjmy těch,

<sup>2</sup> Detailní popis průzkumů včetně dotazníků, souhrnných tabulek a výsledků explicitně zmíněných v textu je k dispozici na vyžádání od autorů na adrese <http://home.cerge-ei.cz/hanousek/evasion>.

TABULKA 1 Hodnoty a 95% intervaly spolehlivosti pro relativní četnosti neplacení daní

## A. Výsledky z průzkumu z roku 2000: roky 1995, 1999, 2000

rok	šedá ekonomika	oficiální ekonomika
1995	15,8 % (13,5 %, 18,1 %)	84,2 % (81,9 %, 86,5 %)
1999	20,3 % (17,8 %, 22,9 %)	79,7 % (77,1 %, 82,2 %)
2000	25,1 % (22,4 %, 27,9 %)	74,9 % (72,1 %, 77,6 %)

## B. Výsledky z průzkumu z roku 2002: roky 1997, 2000, 2002

rok	šedá ekonomika	oficiální ekonomika
1997	23,1 (20,4 %, 25,8 %)	76,9 % (74,2 %, 79,6 %)
2000	25,9 % (23,1 %, 28,7 %)	74,1 % (71,3 %, 76,9 %)
2002	23,9 % (21,2 %, 26,6 %)	76,1 % (73,4 %, 78,9 %)

## C. Výsledky z průzkumu z roku 2004: roky 1999, 2002, 2004

rok	šedá ekonomika	oficiální ekonomika
1999	22,2 % (19,5 %, 24,8 %)	77,8 % (75,2 %, 70,5 %)
2002	23,2 % (20,5 %, 25,8 %)	76,8 % (74,2 %, 79,5 %)
2004	21,4 % (18,7 %, 24,0 %)	78,7 % (76,0 %, 81,3 %)

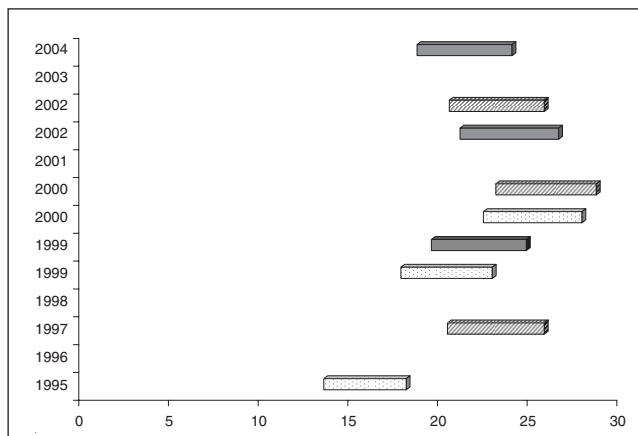
poznámka: Vzorce pro odhad intervalů spolehlivosti jsou uvedeny v Příloze.

zdroj: autorské výpočty založené na průzkumech z let 2000–2004

kteří uvedli, že jsou v šedé ekonomice vysoce aktivní, a těch, kteří v ní působí pouze „občas“, byly od sebe statisticky nerozlišitelné; a stejně tomu bylo i v ostatních demografických charakteristikách těchto dvou skupin. Navíc se tyto dvě kategorie téměř nelišily v rozsahu daňového úniku – u těch, kteří byli ochotni ho prozradit. Z důvodu malých rozdílů mezi těmito dvěma skupinami jsme se rozhodli je sloučit do jedné kategorie, kterou jsme nazvali „šedá ekonomika“. K tomuto kroku jsme se rozhodli také proto, že rozlišování klade příliš vysoké nároky na data a přínos v podobě rozpoznání, proč lidé přecházejí mezi kategoriemi „velmi často“ a „občas“, je pochybný. *Tabulka 1* – části A, B a C – je odvozena z našich průzkumů v letech 2000, 2002 a 2004; ukazuje četnosti odpovědí, že respondenti neplatili daně v různých letech. Pro každou kategorii neplacení jsme vypočítali 95% intervaly spolehlivosti (viz detaily v technické příloze).

Co lze říci o trendech ve vývoji šedé ekonomiky? V *grafu 1* jsme spojili všechny tři průzkumy a složili jsme časovou řadu z bodových a intervalových odhadů neplacení daní. Graf ukazuje vzestupnou tendenci v druhé polovině devadesátých let v počtu těch, kteří přiznávají, že jsou aktivní v šedé

GRAF 1 Trendy v odhadnutých intervalech spolehlivosti pro aktivity v šedé ekonomice



*poznámka:* Pro roky 2002, 2000 a 1999 máme odhady šedé ekonomiky z různých (retrospektivních) šetření; proto jsou tyto roky uvedeny na ose y dvakrát. Vzorčky na proučcích označují odhadnuté intervaly spolehlivosti pro šedou ekonomiku odpovídající průzkumům v roce 2000 (nejsvětlejší), 2002 (střední), 2004 (nejtmavší).

ekonomice. Poté jejich procento mírně klesá. Provedli jsme  $\chi^2$ -kvadrát testy ukazující, že mezi roky 1995 a 1997 šedá ekonomika měřená jako počet neplatičů daní vzrůstala, poté se počet neplatičů ustálil a klesal až do roku 2004.

## 2.2 Konzistence průzkumů

Jedním z cílů tohoto článku je spojit dohromady tři nezávislé průzkumy a provést regrese, které by nám umožnily definovat proměnné určující pravděpodobnosti přechodu mezi jednotlivými kategoriemi placení a neplacení daní. Vzhledem k tomu, že naše průzkumy nejsou panelovými daty, ale nezávislými výzkumy prováděnými ve dvouletých intervalech, nabízí se otázka, zda je spojení těchto průzkumů vůbec možné.

Jednou z hlavních proměnných, které nás při tomto výzkumu zajímají, je odpověď jednotlivce na to, zda platil, nebo neplatil daně. V každém průzkumu jsme se ptali lidí na jejich současné i minulé daňové úniky. Kdyby se nám podařilo zjistit, že odpovědi o neplacení daní v roce 2002 získané při výzkumu provedeném v roce 2004 jsou statisticky nerozeznatelné od odpovědí na stejnou otázku při průzkumu v roce 2002, můžeme usuzovat, že paměť lidí je dobrá a výzkumy v letech 2002 a 2004 jsou navzájem konzistentní v tom smyslu, že průzkumy lze spojit a vytvořit tak časovou řadu dat týkající se šedé ekonomiky. Spojování souborů dat tímto způsobem můžeme zdůvodnit faktem, že v každém průzkumu získáváme retrospektivní data o dvou minulých obdobích. Pokud je paměť lidí dobrá – a naše čísla naznačují že ano –, pak můžeme s jistotou vytvořit časovou řadu, která bude mít pro každý průzkum tři datové body.

Tabulka 1 A–C se zdá ve smyslu předchozího odstavce konzistentní; přes-

to potřebujeme formální test, který by potvrdil náš odhad. Obecně, mějme dva nezávislé náhodné výběry o  $n$  a  $m$  pozorováních,  $\mathbf{x}_1 = (x_{11}, x_{12}, \dots, x_{1n})$  a  $\mathbf{x}_2 = (x_{21}, x_{22}, \dots, x_{2m})$ , kde  $x_{ij}$  označuje  $j$ -té pozorování z  $i$ -tého průzkumu. Průzkum 1 byl proveden v roce 2000 a průzkum 2 v roce 2002. Vektor  $\mathbf{x}_1$  v průzkumu z roku 2000 jsou odpovědi respondentů na otázku, zda neplatili daně v roce 2000, a vektor  $\mathbf{x}_2$  v průzkumu z roku 2002 jsou odpovědi respondentů na otázku, zda neplatili daně rovněž v roce 2000. Odpovědi „ano“ jsou kódovány jako jedničky a odpovědi „ne“ jako nuly. Data nejsou panellová, neboť v každém průzkumu odpovídali jiní respondenti. Proměnné, které nás zajímají, je podíl neplatičů v každém vzorku, tedy:

$$p_1 = \frac{1}{n} \sum_{i=1}^n x_{1i} \text{ a } p_2 = \frac{1}{m} \sum_{i=1}^m x_{2i}$$

Chceme testovat hypotézu:

$$H_0: p_1 = p_2$$

tedy že podíl neplatičů v obou vzorcích je stejný. Uvažujme následující testovou statistiku:

$$u = \frac{p_1 - p_2}{\sqrt{\bar{p}(1 - \bar{p}) \left( \frac{1}{n} + \frac{1}{m} \right)}} \quad (1)$$

kde  $\bar{p} = \frac{1}{(n + m)} (np_1 + mp_2)$ . Pokud platí nulová hypotéza, má testová statistika  $u$  asymptoticky standardní normální rozdělení. Tato testová statistika nám umožní rozlišit, zda byly určité proměnné vybrány z rozdílných rozdělení. Naše výsledky shrnuje *tabulka 2*.

Celkově vede tabulka 2 A–C ke konstatování, že:

1. není možné nalézt žádný rozdíl mezi odhady šedé ekonomiky pro rok 2000 provedenými na základě průzkumů z roku 2000 a z roku 2002, protože  $U$ -statistika  $U = -0,359$  není signifikantní;
2. stejný výsledek dostáváme pro odhady šedé ekonomiky v roce 2002 založené na průzkumech z let 2002 a 2004 ( $U = 0,382$  není signifikantní);
3. stejné tvrzení můžeme uvést i pro odhad šedé ekonomiky v roce 1999 na základě průzkumu z roku 2000 a stejný odhad na základě průzkumu z roku 2002 ( $U = -0,863$ , není signifikantní).

Tím, že jsme prokázali silnou vazbu mezi jednotlivými průzkumy, jsme nejen ospravedlnili možnost sloučení průzkumů, ale zároveň jsme odkryli fakt, že odpovědi na otázky ohledně minulých daňových úniků v průzkumu z jednoho roku jsou statisticky nerozlišitelné od odpovědí na otázky ohledně současných daňových úniků z průzkumu provedeného o dva roky dříve. Přestože tyto dva průzkumy byly provedeny nezávisle na sobě, nabízí se možnost říci, že *lidé si důležité věci pamatují*. Nejsme si úplně jisti, jak tento výsledek zapadá do kontextu současné analýzy, ale považujeme ho za významnou poznámku k našemu výzkumu, která může být zajímavá sama o sobě. Samozřejmě zůstává otázkou, zda jsou si respondenti schopni za-

TABULKA 2 Testy konzistence průzkumů: porovnání retrospektivních odhadů neplacení daní

## A. Šedá ekonomika v roce 2000 (test konzistence 2000 a 2002)

průzkum	šedá ekonomika	oficiální ekonomika	celkem
2000	268	794	1 062
2002	268	766	1 034
testová statistika	<b>-0,359</b>	<i>p</i> -hodnota	<b>0,360</b>

## B. Šedá ekonomika v roce 2002 (test konzistence 2002 a 2004)

průzkum	šedá ekonomika	oficiální ekonomika	celkem
2002	247	788	1 035
2004	245	813	1 058
testová statistika	<b>0,382</b>	<i>p</i> -hodnota	<b>0,649</b>

## C. Šedá ekonomika v roce 1999 (test konzistence 2000 a 2004)

průzkum	šedá ekonomika	oficiální ekonomika	celkem
2000	219	843	1 062
2004	234	822	1 056
testová statistika	<b>-0,863</b>	<i>p</i> -hodnota	<b>0,194</b>

pamatovat své daňové úniky z doby před pěti lety. Naše jediná kontrola paměti respondentů spočívá v konzistenci otázek ohledně neplacení daní ve výzkumech z let 2000, 2002 a 2004. Bystrý kritik může namítnout, že pokud považujeme vzestupný trend v neplacení daní, který je zjevný z průzkumu v roce 2000, za monotónní, měla by míra daňových úniků pro rok 1997 získaná z průzkumu v roce 2002 ležet mezi hodnotami za roku 1995 a 1999. To ovšem znamená, že podle průzkumu z roku 2000 by měla být mezi 15,8 % a 20,3 %, ale podle průzkumu z roku 2002 je průměrná míra neplacení daní v roce 1997 23,1 %. To je jediné místo v našich průzkumech, které naznačuje určité problémy s jednoduchou interpolací v minulých obdobích. Bohužel nemáme retrospektivní dotazování na rok 1997, které by mohlo tento odlišný výsledek podpořit. Rok 1997 byl relativně specifický (měnová krize a „úsporné balíčky“), což mohlo ve svém důsledku vést ke krátkodobému zvýšení šedé ekonomiky tak, jak vyplynulo z odpovědí našich respondentů. Obecně věříme, že retrospektivní odhady daňových úniků mohou být poctivější než odhady současné. Respondent nemůže být snadno potrestán za daňový únik ze vzdálené minulosti. Jeho motiv lhát ohledně minulých daňových úniků je slabší než motiv nemluvit pravdu o svém současném neplacení daní. Pak bychom měli očekávat, že retrospektivní odpovědi v roce 2002 ohledně daňových úniků v roce 1997 budou vyšší než to, co respondenti v roce 2000 tvrdili o nedávné minulosti. To může být důvodem toho, proč je bodový odhad procenta daňových úniků pro rok 1997 vyšší, pokud použijeme průzkum z roku 2002, než když interpolujeme mezi hodnotami pro roky 1995 a 1999 v průzkumu z roku 2000.



TABULKA 3 Markovská matice přechodu mezi roky 1995 a 2000

1995/2000		2000	
		šedá ekonomika	oficiální ekonomika
1995	šedá ekonomika	$P_{nn}$	$P_{np}$
	oficiální ekonomika	$P_{pn}$	$P_{pp}$

### 3. Jak se vyvíjí šedá ekonomika

Nyní se soustředíme na pravděpodobnost přechodu mezi dvěma stavy, placením a neplacením daní, což nám umožní odhadnout budoucí vývoj šedé ekonomiky v České republice. Při svém vstupu do pracovního procesu začne určitá část pracovníků platit daně, jiná však ne. Ti, kteří daně platí od zahájení pracovní činnosti, mohou jednoho dne přestat, a naopak ti, kteří jsou neplatiči, se mohou napravit. Tyto přesuny mezi šedou a oficiální ekonomikou shrnujeme v  $2 \times 2$  matici přechodu pro každého jednotlivce v průzkumu z roku 2000 – viz *tabulka 3*.

Každá buňka udává pravděpodobnost, že jednotlivec přejde z jednoho stavu v roce 1995 do jiného stavu v roce 2000. Například,  $P_{np}$  udává pravděpodobnost, že jednotlivec, který neplatil daně v roce 1995, je bude platit v roce 2000.

Předpokládáme, že přechod z šedé ekonomiky do ekonomiky oficiální (a obráceně) splňuje markovskou vlastnost; to znamená, že nejlepší předpověď budoucích přechodů závisí pouze na současném chování daňového poplatníka. Jinak řečeno: markovská vlastnost stanovuje, že dráha, po které jednotlivec dorazil do současného stavu, neovlivňuje pravděpodobnosti přechodu do budoucna. Neplacení daní ve třech předchozích obdobích nezvyšuje pravděpodobnost dalšího neplacení oproti případu, kdy daný jedinec neplatil daně pouze v posledním období. Pro ty, kteří zastávají teorii „učit se praxí“, může být markovská vlastnost nežádoucí. My ale nemodelujeme závislost současných pravděpodobností přechodu na minulém chování částečně z praktických důvodů. K předpovědi pomocí závislosti na celé minulosti je třeba znát vliv minulých stavů a vliv individuálních vlastností jedince na určení přechodu z jednoho stavu do druhého. My však nemáme dostatečně dlouhou časovou řadu, abychom tyto dvě složky mohli rozlišit, ani nevíme, jak by toto opomenutí mohlo nebo nemohlo ovlivnit naše předpovědi. Vzhledem k tomu, že jde o nový přístup v současném výzkumu, chceme odhalit, co bude třeba udělat v budoucnu, a zároveň vysvětlit, že i když může markovská vlastnost vypadat příliš silná, nepřijali jsme tento předpoklad bezdůvodně.

Pravděpodobnost přechodu mezi kategoriemi šedé a oficiální ekonomiky je pro každého jednotlivce individuální a bude se lišit od pravděpodobností přechodu ostatních osob. K přesnému odhadu budoucího vývoje neplacení daní bychom potřebovali spočítat matici přechodu pro každého jednotlivce a zjistit procentní podíl přechodů jednotlivce z buňky do buňky. Potom by-



chom sečetli tyto pravděpodobnosti v každém roce, a tím získali celkový počet neplatců v každé ze dvou kategorií. Jednodušeji, i když s trochu nižší přesností, lze získat stejný výsledek výpočtem celkových pravděpodobností přechodu. To lze provést výpočtem procentního podílu těch, kteří se přesunuli z buňky do buňky mezi roky 1995 a 2000. Tyto celkové pravděpodobnosti budou méně přesné, než kdybychom použili matici přechodu pro každého jedince, ale díky velkému počtu pozorování (zákon velkých čísel) rozptýl našich výpočtů kolem skutečné střední hodnoty (za předpokladu navzájem nezávislých individuálních pravděpodobností přechodu) nebude příliš vzdálený od své skutečné hodnoty.

Naše technika pro predikci vývoje šedé ekonomiky je nová a je třeba ji chápat v kontextu předchozího výzkumu v této oblasti. Práce (Allingham – Sandmo, 1972), (Watson, 1985), (Jung et al., 1994), (Yaniv, 1994) a další pokládají neplacení daní za rizikové rozhodnutí. Poplatník porovnává riziko odhalení se zisky z neplacení daní. Takové modely se zabývají zejména optimálními politikami kontroly a odhalování daňových úniků a nacházíme je v literatuře o ekonomii zločinu. Nesnaží se modelovat vývoj šedé ekonomiky za delší období.

Engle a Hines (1999) navazují na předchozí modely simulací a testováním modelu dlouhodobé dynamiky šedé ekonomiky založeného na agregátních datech ze Spojených států. Lidé mimo daňový úřad nevědí přesně, jak rozhodování o daňové kontrole u poplatníka probíhá. Předpokládejme, že pravděpodobnost daňové kontroly je rostoucí funkcí daňových úniků poplatníka. Engle a Hines (1999) zabudovali tuto hypotézu do svého modelu, ve kterém je poplatníkům současný daňový únik klesající funkcí minulých úniků, „protože pokud byl kontrolován a přistižen při daňových únicích letos, může být penalizován i za minulé úniky“. Agregátní daňové úniky vykazují cyklický pohyb, když dostatečné množství poplatníků „cykluje“ synchronizovaně, což se stává pod vlivem agregátních šoků, které často ovlivňují všechny poplatníky stejně. Engle a Hines objevili zajímavý výsledek: Pokud nenastanou agregátní šoky, konverguje průměrná hodnota neplacení daní ke stálému stavu a agregátní daňové úniky se blíží limitní hodnotě, přestože hodnoty pro jednotlivce stále vykazují cyklické chování. Tato odlišnost mezi celkovými a individuálními cykly vzniká proto, že chování jednotlivce závisí na tom, zda byl, či nebyl kontrolován, zatímco stav celé ekonomiky je výsledkem rozdělení jednotlivých kontrol mezi poplatníky s různou historií neplacení. Rozdíl mezi celkovými a individuálními cykly v neplacení daní je podobný jako rozdíl mezi poměrem pohlaví v rodině a ve společnosti.

Ze závěru, ke kterému dospěli Engle a Hines (1999), totiž že velikost šedé ekonomiky konverguje ke stálému stavu, můžeme dovodit, jak se bude vyvíjet neplacení daní v České republice. Engle a Hines použili svůj model ke zkoumání spojených agregátních dat o daňových únicích. Naše data jsou individuální, diskrétní a mají rozpětí pěti let, během nichž můžeme vidět, zda respondent přeskočil mezi kategoriemi placení a neplacení daní. Podle našich znalostí je takový soubor dat unikátní.

K pochopení hlavních rysů našeho souboru dat a toho, jak nám umožňuje předpovídat budoucí velikost šedé ekonomiky, je třeba si uvědomit, že jsme schopni odhadovat jak krátkodobé, tak dlouhodobé markovské

TABULKA 4 Krátkodobá matice přechodu pro roky 1999 a 2000

1999/2000		2000		
		šedá ekonomika	oficiální ekonomika	celkem
1999	šedá ekonomika	216 100,0 % 20,3 %	0 0 0,0 %	216  20,3 %
	oficiální ekonomika	51 6,0 % 4,8%	795 94,0 % 74,9 %	846  79,7 %
	celkem	267 25,1 %	795 74,9 %	1 062

poznámka: Zbývající krátkodobé matice přechodu jsou k dispozici na <http://home.cerge-ei.cz/hanousek/evasion>, případně na vyžádání.

TABULKA 5 Dlouhodobá matice přechodu pro roky 1995 a 2000

1995/2000		2000		
		šedá ekonomika	oficiální ekonomika	celkem
1995	šedá ekonomika	168 100,0 % 15,8 %	0 0,0 % 0,0 %	168  15,8 %
	oficiální ekonomika	100 11,2 % 9,4%	795 88,8 % 74,8 %	895  84,2 %
	celkem	268 25,2 %	795 74,8 %	1 063

poznámka: Zbývající dlouhodobé matice přechodu jsou k dispozici na <http://home.cerge-ei.cz/hanousek/evasion>, případně na vyžádání.

pravděpodobnosti přechodu mezi šedou a oficiální ekonomikou. Krátkodobé pravděpodobnosti můžeme odhadnout z matice uvedené v *tabulce 4*.

Při čtení tabulky 4 je třeba si uvědomit, co znamená hodnota 100 % v buňce vlevo nahoře. Tato buňka ukazuje, že 100 % těch, kdo byli aktivní v šedé ekonomice v roce 1999, byli rovněž aktivní v tomto segmentu i v roce 2000, a toto procento představuje 216 pozorování, tedy 20,3 % ze všech daňových poplatníků v roce 1999. Tabulku 4 nazýváme krátkodobou maticí přechodu, protože počítá markovské pravděpodobnosti přechodu získané z přechodů uskutečněných mezi nejbližším předcházejícím obdobím před dotazováním a tehdejší současností. Dlouhodobá matice přechodu sleduje přechody mezi nejvzdálenějším obdobím a současností v době průzkumu. *Tabulka 5* ukazuje takovou dlouhodobou matici pravděpodobností přechodu pro průzkum z roku 2000 a pro přechody mezi roky 1995 a 2000.

*Tabulky A2a–c a A3a–c v Příloze* ukazují šest markovských matic přechodu: dvě pro každý průzkum, z nichž jedna je krátkodobá a druhá dlouhodobá. Pomocí matic přechodu můžeme vytvořit předpověď pro blízkou i vzdálenější budoucnost. Přesnost předpovědi samozřejmě klesá s rostou-

TABULKA 6 Předpovědi s použitím pevných markovských (dlouhodobých) matic přechodu

rok	šedá ekonomika			oficiální ekonomika		
	2000	2002		2000	2002	
2000	25,1 %		2000	25,1 %		2000
2001	26,9 %		2001	26,9 %		2001
2002	28,6 %	25,9 %	2002	28,6 %	25,9 %	2002
2003	30,2 %	26,3 %	2003	30,2 %	26,3 %	2003
2004	31,9 %	26,7 %	2004	31,9 %	26,7 %	2004
2005	33,6 %	27,1 %	2005	33,6 %	27,1 %	2005
2006	35,1 %	27,5 %	2006	35,1 %	27,5 %	2006
2007	36,5 %	27,9 %	2007	36,5 %	27,9 %	2007
2008	38,0 %	28,2 %	2008	38,0 %	28,2 %	2008
2009	39,5 %	28,5 %	2009	39,5 %	28,5 %	2009
2010	41,0 %	28,8 %	2010	41,0 %	28,8 %	2010

*poznámka:* Tučná čísla jsme odhadli přímo z průzkumu; odhadnuté pravděpodobnosti byly založeny na pětileté matici přechodu. Čísla psaná kurzívou odpovídají lineární interpolaci mezi tučně vyznačenými hodnotami.

cím časovým rozpětím. *Tabulka 6* ukazuje naše předpovědi při použití dlouhodobých pravděpodobností přechodu.

První tři sloupce pod nadpisem *šedá ekonomika* ukazují předpovědi vývoje šedé ekonomiky v České republice do roku 2010<sup>3</sup> za použití průzkumů z let 2000, 2002 a 2004. Předpovědi na základě průzkumu z roku 2000, z něhož jsme získali data retrospektivně do roku 1995, ukazují, že šedá ekonomika by se měla téměř zdvojnásobit. Naproti tomu průzkumy z let 2002 a 2004 naznačují buď pokles, nebo vyrovnání velikosti. Rozdíl lze připisat boji proti daňovým únikům vyhlášenému českou vládou v roce 2002. Přesto, že se pravděpodobnosti přechodu změnily jen málo oproti původním odhadům podle roku 2000, vede tato malá změna k velkým změnám v budoucí velikosti šedé ekonomiky. Daný výsledek lze politické veřejnosti prezentovat podobně jako složené úročení vkladů: malá změna dnes může mít značné následky později.

Zjistili jsme, že dlouhodobé i krátkodobé pravděpodobnosti přechodu vytvářejí předpovědi stejného směru, dlouhodobé předpovědi uvedené v tabulce 6 jsou však méně radikální než při použití krátkodobé matice přechodu. Výsledek, že při použití dlouhodobé matice přechodu jsou předpovědi méně výrazné než při použití krátkodobé matice, může být dílem náhody. V delším období se mohou chyby nahromadit nebo naopak vyrušit. Nicméně krátkodobé a dlouhodobé matice přechodu stanovují hranice, mezi kterými se budou pohybovat skutečné tendence vývoje neplacení daní.

<sup>3</sup> Označme  $Pr[p(t)]$  a  $Pr[n(t)]$  po řadě pravděpodobnosti placení daní a neplacení daní v čase  $t$ . Pravděpodobnosti přechodu  $Pr(p(t), p(t+5))$  a  $Pr(n(t), e(t+5))$  můžeme použít k odhadu „budoucích“ pravděpodobností daňových úniků s využitím vzorce

$$Pr[n(t+5)] = Pr[n(t)] \cdot Pr[n(t), n(t+5)] + Pr[p(t)] \cdot Pr[p(t), n(t+5)]$$

Pokud použijeme dlouhodobou matici přechodu, dostaneme samozřejmě předpověď pravděpodobností pouze pro každý pátý rok. Úroveň neplacení daní mezi těmito roky lze dopočítat lineární interpolací.

### 3.1 Stabilita markovských pravděpodobností

Při výpočtech, které zde neuvádíme, jsme zjistili, že při použití kteréhokoli průzkumu podíl šedé a oficiální ekonomiky konverguje ke konstantnímu číslu. Tato konvergence je podobná objevu stálého stavu při daňových únicích, který uskutečnili Engle a Hines (1999). Naše práce jde ještě o krok dále než práce těchto dvou autorů v tom, že k získání předpovědi používáme pravděpodobnosti přechodu spočítané z individuálních dat. Musíme být proto připraveni čelit kritice, že naše konvergence ke stálému stavu je pouze důsledkem použité výpočetní metody. Podstatou našich výpočtů je předpoklad stabilních pravděpodobností přechodu. V transformující se zemi může být odůvodnitelné tvrdit, že tyto pravděpodobnosti nebudou stabilní a v důsledku toho se bude měnit šedá ekonomika spíše cyklicky než konvergovat ke stálému stavu. Tato záležitost nemůže být rozhodnuta teoreticky, ale je nutné ji vyřešit na základě faktů.

Nemáme žádný důvod se domnívat, že pravděpodobnosti přechodu jsou stabilní. Naopak, vzhledem k měnícímu se politickému klimatu a demografickému vývoji České republiky v průběhu studovaného období by nás spíše překvapilo zjištění, že tyto pravděpodobnosti stabilní jsou. Stabilní pravděpodobnosti přechodu by byly žádoucí pro předpovědi vývoje šedé ekonomiky, ale nejsou nezbytnou záležitostí pro odhad faktorů ovlivňujících pravděpodobnosti přechodu ani pro předpověď, jak se budou tyto pravděpodobnosti měnit v závislosti na různých scénářích změn proměnných, které je určují.

Pro předpovědi pravděpodobností přechodu bude třeba odhadnout strukturní tvar koeficientů pro politické a demografické faktory, které je ovlivňují. Než to však učiníme, pokusíme se ověřit stabilitu pravděpodobností přechodu. K prověření stálosti těchto pravděpodobností můžeme využít  $\chi^2$ -kvadrát test. Ten ověřuje, zda můžeme tvrdit, že pravděpodobnosti vypočítané na základě každého ze tří průzkumů pocházejí z výběrů ze stejného rozdělení (obecné vytvoření testu je popsáno v *Příloze*).

Pro tento test stability (homogenity) markovských matic přechodu, dlouhodobých i krátkodobých, uvažujeme následující čtyři kategorie:  $[N \rightarrow N]$  (ten, kdo neplatil daně v předchozím období a pokračuje v tom v současné době),  $[N \rightarrow P]$  (ten, kdo neplatil daně v předchozím období a nyní je platit začal),  $[P \rightarrow N]$  (ten, kdo platil daně v předchozím období a nyní je platit přestal) a  $[P \rightarrow P]$  (ten, kdo platil daně v předchozím období a platí je i nyní). Počty  $n$  přecházejících z jednoho stavu do druhého jsou základem pro výpočet našich pravděpodobností přechodu. Použitím standardního testu homogenity distribucí můžeme neparametricky testovat, zda jsou matice přechodu, které byly použity jako podklad pro výpočty, shodné pro všechny průzkumy. Výsledky shrnujeme v kontingenční *tabulce* 7.

Za předpokladu platnosti nulové hypotézy má testová statistika následující tvar:

$$\chi^2 = \sum_{i=1}^3 \sum_{j=1}^4 \frac{\left(n_{ij} - \frac{1}{n} n_{i.} n_{.j}\right)^2}{\frac{1}{n} n_{i.} n_{.j}} = n \cdot \sum_{i=1}^3 \sum_{j=1}^4 \frac{n_{ij}^2}{n_{i.} n_{.j}} - n \quad (2)$$

a je asymptoticky rozdělená jako  $\chi^2$ -kvadrát se 6 stupni volnosti.

TABULKA 7 Kontingenční tabulka pro testování stability dlouhodobých matic přechodu

výběr (průzkum)	změna v kategorii placení daní				celkem
	1 [ $N > M$ ]	2 [ $N > P$ ]	3 [ $P > M$ ]	4 [ $P > P$ ]	
1 (2000)	$n_{11}$ 168	$n_{12}$ 0	$n_{13}$ 100	$n_{14}$ 795	$n_1$ 1 063
2 (2002)	$n_{21}$ 194	$n_{22}$ 44	$n_{23}$ 73	$n_{24}$ 720	$n_2$ 1 031
3 (2004)	$n_{31}$ 148	$n_{32}$ 78	$n_{33}$ 86	$n_{34}$ 741	$n_3$ 1 053
celkem	$n_{.1}$ 510	$n_{.2}$ 122	$n_{.3}$ 259	$n_{.4}$ 2 256	$N$ 3 147

Hodnota testové statistiky 88,54 naznačuje, že pokud vzájemně porovnáme pravděpodobnosti přechodu, nelze naše tři průzkumy považovat za výběry ze stejného rozdělení.<sup>4</sup> Pravděpodobnosti přechodu by se tedy mohly měnit od průzkumu k průzkumu. Tento fakt může na první pohled vrhnout stín pochybnosti na naši snahu vytvořit rámec pro předpovídání šedé ekonomiky. My ale tento výsledek interpretujeme optimisticky. V transformující se ekonomice by bylo zklamáním, kdybychom nenašli měnící se pravděpodobnosti přechodu.<sup>5</sup> Vyřešit to lze zjištěním toho, které faktory ovlivňují pravděpodobnosti přechodu, a odhadem budoucích hodnot těchto faktorů. Potom můžeme tento odhad využít v rámci markovského modelu k vytvoření předpovědí. Pomocí regresí stanovíme sílu, s jakou jednotlivé náhodné proměnné ovlivňují pravděpodobnosti přechodu. Tyto síly jsou regresní koeficienty. Když budeme mít určeny tyto koeficienty, můžeme simulovat, jak různé hodnoty nezávislých proměnných ovlivní markovské pravděpodobnosti přechodu. Následně upravíme naše odhady budoucích pravděpodobností přechodu v závislosti na tom, jak se tyto nezávislé proměnné budou měnit.

#### 4. Jak zmírnit daňové úniky

Hlavní problém při vytváření spolehlivých odhadů vývoje šedé ekonomiky jsou stále změny demografického a politického rázu. Jak populace bohatne, mění se její přístup k neplacení daní. Když začne vláda bojovat proti daňovým únikům nebo když změní kvalitu služeb poskytovaných občanům, lidé se budou vždy znovu rozhodovat, zda daně platit, nebo neplatit. Tyto změny jsou podstatou Lucasovy kritiky a prokletím všech prognostiků. Když se agregátně změní parametry tvořící základ rozhodnutí neplatit daně, musí se změnit také námi spočítané celkové markovské pravděpodobnosti. Už

<sup>4</sup> Podobně má testová statistika pro stabilitu krátkodobých matic přechodu hodnotu 149,4 a ukazuje, že krátkodobé pravděpodobnosti přechodu také nemohou být pokládány za stabilní.

<sup>5</sup> Poznamenejme, že pokud podobné testy stability provádíme pro proměnné, které mohou být spojeny se změnou v kategorii šedé a oficiální ekonomiky, jako je příjem domácnosti (v porovnání s jeho hodnotami před rokem a před pěti lety), spokojenost s vládními službami atd., také u těchto proměnných stabilitu zamítáme.

jsme ukázali, že se nedomníváme, že jsou tyto pravděpodobnosti v České republice stabilní, přestože nestabilní pravděpodobnosti zpochybňují naše předpovědi ohledně míry neplacení daní. To nejlepší, co můžeme udělat pro obnovení důvěry v naše odhady, je, že odhady pozměníme tím, že odhadneme budoucí vývoj parametrů ovlivňujících šedou ekonomiku a upravíme námi spočítané celkové markovské pravděpodobnosti. Technicky řečeno, chceme regresi získat odhad vlivu rozhodujících faktorů na *přechod* mezi placením a neplacením. Pokud budeme mít odhady vlivu těchto faktorů, můžeme prohlásit, že pokud se bude demografický nebo politický vývoj ubírat určitým směrem, budou se markovské pravděpodobnosti přechodu ubírat také určitým směrem. Na základě znalosti směřování markovských pravděpodobností přechodu pak můžeme upravit své odhady i pro měnící se budoucnost.

Je hezké říci, že chceme odhadnout vliv jistých proměnných na pravděpodobnosti přechodu, ale otázkou zůstává, na jaké proměnné bychom se měli zaměřit? Tato otázka je základem mnoha nedokonalostí v současném přístupu k šedé ekonomice. Prognostici rádi sledují redukovaný tvar odhadu koeficientů u proměnných, které nevstupují do vzájemného vztahu se závislou proměnnou. I my můžeme odhadnout redukovaný tvar regrese určující markovské pravděpodobnosti přechodu, ale takové odhady nám neodhalí, které politiky jsou významné pro ovlivnění těchto pravděpodobností. Politické proměnné do našeho odhadu pravděpodobností přechodu nemůžeme zahrnout jednoduše proto, že neznáme žádný objektivní způsob, jak měřit změny politické strategie bez dlouhodobých časových řad a využití případových studií. Můžeme se respondentů nanejvýš zeptat, jak vnímají vládní politiku; jenže dojem je příliš vrtkavá záležitost na to, abychom ho mohli zařadit do regrese. Nemůžeme také vědět, zda respondenti neuvedli horší pocit z vládní politiky proto, aby si zdůvodnili vlastní neplacení daní: zeptejte se mě, jestli neplatím daně, a já řeknu ano, a potom se mě zeptejte, jestli je morální neplatit. Řeknu rovněž ano, abych vypadal dobře v očích tazatele (rozhovory jsou prováděny osobně). Moje odpovědi budou potom mást výzkumníka, který odhaduje rovnice v redukovaném tvaru, a budou ho nutit k odhadu rekurzivního nebo simultánního modelu daňových úniků. Široce akceptovaný rovnovážný model daňových úniků však stále ještě nebyl vytvořen.

V dřívější studii (Hanousek – Palda, 2004) je uvedena metoda, která se částečně vyhýbá komplikovanému modelování strukturních parametrů, ale přitom stále umožňuje zařadit do redukovaného tvaru modelu proměnné typu vnímání vládní politiky. V tomto článku je provedena regrese daňových úniků jednotlivých poplatníků na individuální vnímání kvality vládních služeb. Tato myšlenka, založená na Downsově hlasovací teorii, vychází z toho, že se lidé nedopouštějí daňových úniků jenom ze zjištěných důvodů (více peněz do vlastní kapsy), ale také z morálních pohnutek ("když nedostávám kvalitní vládní služby, budu protestovat tím, že nezaplatím daně"). Aby tazatelé zajistili, že respondenti nebudou ospravedlňovat své daňové úniky odpověďmi o špatných vládních službách, bylo jim řečeno, že se jedná o výzkum kvality (vládních) služeb. Tazatelé pokládali otázky o kvalitě na začátku rozhovoru. Mnohem později přišly na řadu otázky, týkající se placení či neplacení daní. Myslíme si, že pořadí kladených otázek snižuje ne-



žádoucí korelaci mezi odpověďmi na ně; opačné pořadí by dalo respondentům možnost se dodatečně „ospravedlnit“ odpovědí, že neplatí daně proto, že si myslí, že vládní služby mají nízkou kvalitu (další diskuze a výsledky vztahující se k tomuto konkrétnímu fenoménu viz (Hanousek – Palda, 2004)).

Dalšími politickými proměnnými, které teoretická i empirická literatura považuje za významné, jsou vnímaná pravděpodobnost přistižení při daňovém podvodu a vnímaná sankce za neplacení daní. Tyto proměnné zřetelně patří do strukturní regrese. Naše regrese by tedy měla být chápána jako kvaziredukováný tvar, který obsahuje jednoznačně exogenní proměnné jako například demografické jevy, a vnímané politické proměnné, o jejichž exogenitě mohou být pochybnosti.

To, že máme čtyři pravděpodobnosti přechodu ( $N \rightarrow P$ ,  $N \rightarrow N$ ,  $P \rightarrow N$ ,  $P \rightarrow P$ ), naznačuje, že bychom mohli odhadnout rovnice v redukováném tvaru pro každou ze čtyř možných pravděpodobností přechodu zvlášť. Když odhadneme parametry spojené s proměnnými ovlivňujícími šedou ekonomiku, můžeme simulovat, jak se markovské pravděpodobnosti přechodu změní při změně nezávisle proměnných.<sup>6</sup>

*Tabulka 8* ukazuje regresi jedné pravděpodobnosti přechodu, a to přechodu od placení k neplacení daní, v redukováném tvaru. Je mnoho možných kandidátů na proměnné, které ovlivňují pravděpodobnosti přechodu. K zařazení do našich rovnic však musíme vybrat pouze ty nejpravděpodobnější z nich, protože metoda maximální věrohodnosti je technika, jejíž náročnost na data roste exponenciálně s růstem počtu odhadovaných proměnných. Demografické proměnné jako věk a pohlaví jsou standardními zástupnými proměnnými za vektor individuálních charakteristik. Zařadíme také několik regionálních proměnných, jako je velikost města, a konečně individuální vnímání morálnosti daňových úniků, pravděpodobnost přistižení a fakt, že se ekonomická situace jednotlivce zhoršuje. V tabulce 8 vyhledáváme faktory určující daňové úniky. Je mezi nimi významná změna ekonomické situace jednotlivce (zhoršující se ekonomická situace zvyšuje tendenci neplatit daně), zkušenost jednotlivce s nákupem statků, z nichž nebyla zaplacená daň, a vnímaná pravděpodobnost přistižení při páchání daňových úniků.

Co vypovídá *tabulka 8* o stabilitě pravděpodobností přechodu, které používáme k předpovědi vývoje šedé ekonomiky? Jak populace stárne, můžeme předpokládat, že pravděpodobnost přechodu od placení k neplacení daní bude klesat. Nedostatek místa nám brání v podrobnějším popisu, ale je jednoduché představit si různé rychlosti stárnutí populace, zapojit je do *tabulky 8* a přičíst nebo odečíst změnu v pravděpodobnostech přechodu od pravděpodobností, které jsme používali pro předchozí odhady budoucího vývoje neplacení daní. *Tabulka 8* vypovídá rovněž o tom, že pokud se vládě

<sup>6</sup> Čtenáři se mohou ptát, jak naše výpočty ovlivní nově nastupující daňoví poplatníci. Naše data nám neumožňují poznat, kdo je novým poplatníkem. Když předpokládáme, že vstup a výstup z pracovního trhu jsou navzájem stabilní a že jsou rovnoměrně rozděleny mezi jednotlivé kategorie šedé a oficiální ekonomiky, pak do našich výpočtů toho, jak se budou daňové úniky vyvíjet, vstup a výstup z pracovního trhu explicitně zahrnout nemusíme. Nepřímou podporu pro tuto hypotézu dává i náš průzkum, který ukazuje, že ti, kdo na daních podvádějí často, a ti, kdo podvádějí jen občas, mají statisticky nerozlišitelné průměrné příjmy.



TABULKA 8 Výsledky logitové regrese pro Ppn (pravděpodobnost přechodu od placení k neplacení daní) v České republice, marginální vlivy pro spojení všech průzkumů

proměnné	derivace $dP/dX$ přechodu od placení k neplacení daní	
	dlouhodobá	krátkodobá
<i>demografické</i>		
věk	−0,006*	−0,002**
věk (druhá mocnina)	−4E-05	1E-05
žena	−0,035**	−0,013
<i>vzdělání</i>		
základní vzdělání	0,060**	0,023
vyučený	0,032	0,005
vyučený s maturitou	0,028	0,007
střední odborné s maturitou	0,015	−0,002
<i>příjem</i>		
< 10 000	0,014	0,032
10 001 až 15 000	0,039	0,034
15 001 až 20 000	0,023	0,035
20 001 až 25 000	0,064*	0,067
25 001 až 30 000	0,025	0,022**
<i>příjem v porovnání s minulostí</i>		
mnohem horší než před 5 lety	0,071**	0,039
mnohem lepší než před 5 lety	−0,036	−0,006
mnohem lepší než před rokem	−0,001	0,004
<i>demografické pomocné proměnné</i>		
velké město	−0,021	0,003
vesnice	−0,009	−0,019
Praha	0,051**	−0,003
střední Čechy	0,004	0,004
jižní Čechy	0,021	−0,008
západní Čechy	0,047*	0,018
severní Čechy	0,024	0,007
východní Čechy	0,041	0,032**
jižní Morava	0,030**	0,012
<i>faktory spojené se stavem daňových úniků</i>		
Nakupoval statky v šedé ekonomice.	0,036**	0,005
Neplacení daní je morální.	0,006	−0,011
Neplacení daní je velmi nemorální.	−0,028*	−0,017*
pravděpodobnost odhalení	0,001**	−9E-05
$R^2$ (2 859 pozorování)	0,08	0,03

poznámka: Značky \* a \*\* znamenají případy, kdy související koeficienty byly signifikantní na 5%, případně 1% hladině významnosti.

podarí přesvědčit lidi, aby si mysleli, že pravděpodobnost přistížení vzrůstá nebo že neplatit daně je nemorální, daňové úniky poklesnou. Samy o sobě jsou tyto poznatky všedním, byť zajímavým, doplňkem empirické práce o daňových únicích. Ale v kontextu předpovědi vývoje šedé ekonomiky nám tyto výsledky ukazují cestu, jak upravit markovské pravděpodobnosti přechodu, a tím zpřesnit naše předpovědi.

Dalším krokem v předpovídání změn v markovských pravděpodobnostech by bylo odhadnout mnohonásobný logitový model, který bude brát všechny čtyři pravděpodobnosti přechodu jako současně určené. *Tabulka 9* prezen-

TABULKA 9 Výsledky mnohonásobné logitové regrese pro všechny pravděpodobnosti přechodu: marginální vlivy  
(derivace pravděpodobností přechodu podle vysvětlujících proměnných)

proměnné	dlouhodobý			krátkodobý		
	$N \Rightarrow N$	$P \Rightarrow N$	$P \Rightarrow P$	$N \Rightarrow N$	$P \Rightarrow N$	$P \Rightarrow P$
<i>demografické</i>						
věk	0,012	-0,007**	-0,010**	0,006	-0,002	-0,005
věk (druhá mocnina)	0,000	0,000**	0,000**	0,000	0,000	0,000
žena	-0,045	-0,033	0,098*	-0,066	-0,012	0,084
<i>vzdělání</i>						
vyučený	0,065	0,002	-0,104**	0,075	-0,007	-0,090*
vyučený s maturitou	0,052	0,001	-0,074*	0,056	-0,004	-0,071*
střední odborné s maturitou	0,002	-0,012	-0,006	0,003	-0,012	-0,002
<i>příjem</i>						
< 10.000	0,037	0,012	-0,055	0,015	0,033	-0,053
10 001 až 15 000	0,071	0,033	-0,130	0,071	0,033	-0,108
15 001 až 20 000	0,079	0,015	-0,134	0,062	0,033	-0,102
20 001 až 25 000	0,114	0,057	-0,166	0,101	0,066	-0,169
25 001 až 30 000	0,139	0,020	-0,166	0,137	0,022	-0,168
<i>příjem v porovnání s minulostí</i>						
mnohem horší než před 5 lety	-0,052	0,071	-0,038	-0,019	0,039	-0,031
mnohem lepší než před 5 lety	0,055	-0,029	-0,032	0,030	-0,003	-0,031
mnohem lepší než před rokem	0,010	0,001	0,009	0,006	0,003	-0,012
<i>demografické pomocné proměnné</i>						
velké město	0,017	-0,020	0,000	-0,006	0,003	0,009
vesnice	-0,006*	-0,007	0,049*	0,004*	-0,018	0,045*
Praha	-0,025	0,051*	-0,016	0,028	-0,003	-0,028
střední Čechy	-0,021	0,006	0,007	-0,023	0,005	0,028
jižní Čechy	0,004	0,025	-0,040	0,034	-0,007	-0,035
západní Čechy	0,024	0,049	-0,078	0,052	0,020	-0,069
severní Čechy	-0,057	0,026	0,027	-0,041	0,007	0,039
východní Čechy	0,023	0,041	-0,074	0,026	0,032	-0,056
jižní Morava	0,012	0,032	-0,024	0,034**	0,013**	-0,010*
<i>faktory spojené se stavem daňových úniků</i>						
Nakupoval statky v šedé ekonomice.	0,100	0,028**	-0,186**	0,121**	0,000**	-0,185**
Neplacení daní je morální	0,066	0,017	-0,087	0,086	-0,005	-0,084
Neplacení daní je velmi nemorální.	-0,149**	-0,019	0,162	-0,149**	-0,014	0,171**
pravděpodobnost odhalení	-0,002**	-0,001	0,002**	-0,002	0,000	0,003**
$R^2$ (2 859 pozorování)	0,29			0,28		

poznámka: Značky \* a \*\* znamenají případy, kdy související koeficienty byly signifikantní na 5%, případně 1% hladině významnosti.

tuje tyto odhady pro tři ze čtyř pravděpodobností přechodu (nemusíme odhadovat čtvrtou regresi, protože je z definice určena zbývajícími).

Mnoho proměnných v mnohonásobném logitovém modelu ztratí svou významnost, ale tato metoda odhadu koeficientů by nám měla poskytnout přesnější odhady, než když byl logit pro každou pravděpodobnost přechodu odhadován nezávisle na ostatních. Nebudeme se uvedeným odhadům dále věnovat, ale ponecháváme je zde jako ukázkou „dalšího kroku“ v odhadování změn v markovských pravděpodobnostech přechodu.

## 5. Závěr

V předloženém článku jsme nabídli metodu odhadující možný vývoj šedé ekonomiky. V několika průzkumech jsme se ptali jednotlivců, zda nepřestali nebo nezačali platit daně nebo zda u nich v tomto ohledu došlo ke změně. Odpovědi na tyto otázky nám umožnily předpovědět, jak by se úroveň šedé ekonomiky (pokud ne hodnota nezaplacených daní) mohla vyvíjet několik let do budoucna. V reakci na možnou kritiku, že pravděpodobnosti přechodu mezi placením a neplacením daní nemusejí být konstantní, jsme předložili formální metodologii testování stability zmíněných pravděpodobností. V nejhorším možném případě, kdy by pravděpodobnosti přechodu nebyly konstantní, jsme ukázali, jak je možné použít mnohonásobný logitový model pro simulaci. Pokud jsou pravděpodobnosti přechodu funkcí proměnných zahrnujících demografický vývoj a vládní politiky, může mnohonásobný logit stanovit koeficienty nezávislých proměnných určujících pravděpodobnosti přechodu. Předpovědi pak mohou být změněny v závislosti na předpokladech vývoje proměnných, které pravděpodobnosti přechodu ovlivňují.

Tento článek nemá normativní zaměření a soustřeďuje se převážně na metodologii předpovídání vývoje šedé ekonomiky za použití výběrových šetření.

# PŘÍLOHA

TABULKA A1.a Popisné statistiky pro průzkum v roce 2000, Česká republika

průzkum 2000	celkem vzorek	nákup nezdaněného zboží/služeb		aktivní zapojení v šedé ekonomice	
		ano	nikdy	ano	nikdy
<b>celkem</b>	<b>1 062</b>	<b>507</b>	<b>555</b>	<b>268</b>	<b>794</b>
<i>pohlaví</i>					
muž	518	50,0 %	50,0 %	34,6 %	65,4 %
žena	544	45,6 %	54,4 %	16,4 %	83,6 %
<i>věk</i>					
18–25	188	46,8 %	53,2 %	30,3 %	69,7 %
26–35	247	52,6 %	47,4 %	25,9 %	74,1 %
36–45	229	52,8 %	47,2 %	31,4 %	68,6 %
46–55	248	50,8 %	49,2 %	22,2 %	77,8 %
56–65	150	28,0 %	72,0 %	13,3 %	86,7 %
<i>vzdělání</i>					
základní	252	47,6 %	52,4 %	29,0 %	71,0 %
střední bez maturity	427	51,5 %	48,5 %	32,6 %	67,4 %
střední s maturitou	295	45,1 %	54,9 %	14,6 %	85,4 %
vyšší	88	38,6 %	61,4 %	14,8 %	85,2 %
<i>příjem v Kč</i>					
< 10 000	566	45,1 %	54,9 %	22,1 %	77,9 %
10 001 až 15 000	266	54,9 %	45,1 %	33,1 %	66,9 %
15 001 až 20 000	91	54,9 %	45,1 %	23,1 %	76,9 %
20 001 až 25 000	22	59,1 %	40,9 %	50,0 %	50,0 %
25 001 až 30 000	9	55,6 %	44,4 %	55,6 %	44,4 %
30 001 až 40 000	1	100,0 %	0,0 %	100,0 %	0,0 %
odmítl odpovědět	107	34,6 %	65,4 %	15,9 %	84,1 %

zdroj: výzkum v roce 2000; autorské výpočty

TABULKA A1.b Popisné statistiky pro průzkum v roce 2002, Česká republika

průzkum 2002	celkem vzorek	nákup nezdaněného zboží/služeb		aktivní zapojení v šedé ekonomice	
		ano	nikdy	ano	nikdy
<b>celkem</b>	<b>1 041</b>	<b>573</b>	<b>464</b>	<b>247</b>	<b>788</b>
<i>pohlaví</i>					
muž	513	58,9 %	41,1 %	29,8 %	70,0 %
žena	528	51,3 %	47,9 %	17,8 %	81,3 %
<i>věk</i>					
18–25	196	56,6 %	42,9 %	29,6 %	69,4 %
26–35	241	58,1 %	41,5 %	26,1 %	73,4 %
36–45	214	55,6 %	44,4 %	26,6 %	72,4 %
46–55	237	54,0 %	45,1 %	18,6 %	81,4 %
56–65	153	49,0 %	51,0 %	16,3 %	83,0 %
<i>vzdělání</i>					
základní	195	54,9 %	44,6 %	25,1 %	74,4 %
střední bez maturity	404	58,4 %	41,3 %	26,0 %	73,5 %
střední s maturitou	339	53,1 %	46,6 %	23,3 %	75,8 %
vyšší	103	48,5 %	50,5 %	13,6 %	86,4 %

*pokračování tabulky na následující straně*

TABULKA A1.b Popisné statistiky pro průzkum v roce 2002, Česká republika (pokračování)

průzkum 2002	celkem vzorek	nákup nezdaněného zboží/služeb		aktivní zapojení v šedé ekonomice	
		ano	nikdy	ano	nikdy
příjem v Kč					
< 10,000	487	55,4 %	43,9 %	19,5 %	79,9 %
10 001 až 15 000	335	51,0 %	49,0 %	24,5 %	74,9 %
15 001 až 20 000	114	59,6 %	40,4 %	30,7 %	68,4 %
20 001 až 25 000	35	71,4 %	28,6 %	42,9 %	57,1 %
25 001 až 30 000	18	72,2 %	27,8 %	50,0 %	50,0 %
30 001 až 40 000	4	75,0 %	25,0 %	25,0 %	75,0 %
40 001 až 50 000	3	100,0 %	0,0 %	33,3 %	66,7 %
>= 50 001	1				
odmítl odpovědět	44	45,5 %	52,3 %	20,5 %	79,5 %

zdroj: výzkum v roce 2002, autorské výpočty

TABULKA A1.c Popisné statistiky pro průzkum v roce 2004, Česká republika

průzkum 2004	celkem vzorek	nákup nezdaněného zboží/služeb		aktivní zapojení v šedé ekonomice	
		ano	nikdy	ano	nikdy
<b>celkem</b>	<b>1 066</b>	<b>568</b>	<b>494</b>	<b>227</b>	<b>836</b>
<i>pohlaví</i>					
muž	535	58,7 %	40,7 %	29,9 %	69,7 %
žena	531	47,8 %	52,0 %	12,6 %	87,2 %
<i>věk</i>					
18–25	183	58,5 %	41,0 %	25,1 %	74,9 %
26–35	270	55,6 %	44,4 %	23,3 %	76,3 %
36–45	237	51,9 %	46,8 %	16,5 %	83,5 %
46–55	269	50,6 %	49,4 %	23,0 %	76,6 %
56–65	181	47,0 %	53,0 %	18,2 %	81,2 %
<i>vzdělání</i>					
základní	159	59,7 %	39,6 %	28,3 %	71,7 %
střední bez maturity	445	60,0 %	39,6 %	24,5 %	75,5 %
střední s maturitou	347	46,7 %	54,5 %	16,4 %	83,0 %
vyšší	115	38,3 %	61,7 %	13,9 %	85,2 %
<i>příjem v Kč</i>					
< 10,000	448	54,2 %	45,5 %	17,6 %	82,4 %
10 001 až 15 000	331	54,7 %	44,7 %	23,9 %	75,2 %
15 001 až 20 000	139	48,2 %	51,8 %	23,7 %	76,3 %
20 001 až 25 000	41	58,5 %	41,5 %	39,0 %	61,0 %
25 001 až 30 000	28	64,3 %	35,7 %	28,6 %	71,4 %
30 001 až 40 000	2	0,0 %	100,0 %	0,0 %	100,0 %
40 001 až 50 000	1	0,0 %	100,0 %	0,0 %	100,0 %
>= 50 001	0				
odmítl odpovědět	76	46,1 %	52,6 %	15,8 %	84,2 %

zdroj: výzkum v roce 2004, autorské výpočty

TABULKA A2.a Krátkodobá matice přechodu pro 1999/2000 založená na průzkumu z roku 2000

1999/2000		2000		
		šedá ekonomika	oficiální ekonomika	celkem
1999	šedá ekonomika	216 100,0 % 20,3 %	0 0 0,0 %	216  20,3 %
	oficiální ekonomika	51 6,0 % 4,8 %	795 94,0 % 74,9 %	846  79,7 %
	celkem	267 25,1 %	795 74,9 %	1 062

TABULKA A2.b Krátkodobá matice přechodu pro 2000/2002 založená na průzkumu z roku 2002

2000/2002		2002		
		šedá ekonomika	oficiální ekonomika	celkem
2000	šedá ekonomika	218 81,3 % 21,1 %	50 18,7 % 4,8 %	268  25,9 %
	oficiální ekonomika	29 3,8 % 2,8 %	737 96,2 % 71,3 %	766  74,1 %
	celkem	247 23,9 %	787 76,1 %	1 034

TABULKA A2.c Krátkodobá matice přechodu pro 2002/2004 založená na průzkumu z roku 2004

2002/2004		2004		
		šedá ekonomika	oficiální ekonomika	celkem
2002	šedá ekonomika	195 85,9 % 18,5 %	32 14,1 % 3,0 %	227  21,5 %
	oficiální ekonomika	50 6,0 % 4,7 %	778 94,0 % 73,7 %	828  78,5 %
	celkem	247 23,4 %	787 74,6 %	1 055

TABULKA A3.a Dlouhodobá matice přechodu pro 1995/2000 založená na průzkumu z roku 2000

1995/2000		2000		
		šedá ekonomika	oficiální ekonomika	celkem
1995	šedá ekonomika	168 100,0 % 15,8 %	0 0,0 % 0,0 %	168  15,8 %
	oficiální ekonomika	100 11,2 % 9,4 %	795 88,8 % 74,8 %	895  84,2 %
	celkem	268 25,2 %	795 74,8 %	1 063

TABULKA A3.b Dlouhodobá matice přechodu pro 1997/2002 založená na průzkumu z roku 2002

1997/2002		2002		
		šedá ekonomika	oficiální ekonomika	celkem
1997	šedá ekonomika	194 81,5 % 18,8 %	44 18,5 % 4,3 %	238  23,1 %
	oficiální ekonomika	73 9,2 % 7,1 %	720 90,8 % 69,8 %	793  76,9 %
	celkem	267 25,9 %	764 74,1 %	1 031

TABULKA A3.c Dlouhodobá matice přechodu pro 1999/2004 založená na průzkumu z roku 2004

1999/2004		2004		
		šedá ekonomika	oficiální ekonomika	celkem
1999	šedá ekonomika	148 65,5 % 14,1 %	78 34,5 % 7,4 %	226  21,5 %
	oficiální ekonomika	86 10,4 % 8,2 %	741 89,6 % 70,4 %	827  78,5 %
	celkem	267 25,4 %	764 72,6 %	1 053

### Výpočet intervalů spolehlivosti pro odhadnuté pravděpodobnosti daňových úniků

Výběrové relativní četnosti  $p_p$  a  $p_n$  nám umožní sestavit intervaly spolehlivosti pro podkladové pravděpodobnosti  $P_p$  a  $P_n$  a pro celou matici přechodu  $\Pi_{ij}$ . Protože zkoumáme náhodný vzorek z české populace o velikosti několika milionů, můžeme použít známou aproximaci normálním rozdělením (viz například (Cochran, 1963)) a ukázat, že:



$$\frac{p - P}{\sqrt{\frac{p(1-p)}{n-1} \frac{N-n}{N}}} \approx \frac{p - P}{\sqrt{\frac{p(1-p)}{n-1}}} \sim N(0,01) \quad (\text{A.1})$$

kde  $p$  a  $P$  mohou odkazovat popořadě na  $p_{\omega}$  nebo  $p_n$  a  $P_{\omega}$  nebo  $P_n$ . Z toho plyne, že  $(1 - \alpha)\%$  interval spolehlivosti je stanoven jako (použijeme opravu  $1/2n$  pro diskrétní náhodnou veličinu):

$$p - \frac{1}{2n} - u_{1-\alpha/2} \sqrt{\frac{p(1-p)}{n-1}} < P < p + \frac{1}{2n} + u_{1-\alpha/2} \sqrt{\frac{p(1-p)}{n-1}} \quad (\text{A.2})$$

kde  $u$  označuje kvantil standardního normálního rozdělení.

### Test homogenity pro výběrová rozdělení

Předpokládejme, že máme  $I \geq 2$  nezávislých náhodných výběrů o velikostech  $n_1, n_2, \dots, n_I$ . Předpokládejme, že zkoumaná proměnná nabývá hodnot, které se dají rozdělit do  $J$  disjunktních kategorií. Například, máme-li dvě kategorie placení daní (platí, neplatí), pak  $J$  se rovná dvěma. Označme  $n_{ij}$  celkový počet hodnot proměnné v  $i$ -tém vzorku, které spadají do kategorie  $j$ , kde  $i = 1, \dots, I$ , a  $j = 1, \dots, J$ . Když například označíme průzkum z roku 2002 číslem 2, potom  $n_{21}$  bude 247, tj. počet neplatců z tabulky 1b v této příloze. Pro zjištění, zda jsou naše markovské pravděpodobnosti přechodu stabilní, testujeme hypotézu  $H_0$ : všechny výběry  $i = 1, \dots, I$  jsou ze stejného rozdělení. Tato situace je známá jako homogenita podkladových rozdělení  $i = 1, \dots, I$  a dá se shrnout v podobě následující kontingenční tabulky:

výběr	kategorie						celkem
	1	2	...	$j$	...	$J$	
1	$n_{11}$	$n_{12}$	...	$n_{1j}$	...	$N_{1J}$	$n_{1.}$
2	$n_{21}$	$n_{22}$	...	$n_{2j}$	...	$N_{2J}$	$n_{2.}$
...	...	...	...	...	...	...	...
$i$	$n_{i1}$	$n_{i2}$	...	$n_{ij}$	...	$n_{iJ}$	$n_{i.}$
...	...	...	...	...	...	...	...
$I$	$n_{I1}$	$n_{I2}$	...	$n_{Ij}$	...	$n_{IJ}$	$n_{I.}$
celkem	$n_{.1}$	$n_{.2}$	...	$n_{.j}$	...	$n_{.J}$	$N$

Za předpokladu nulové hypotézy má testová statistika

$$\chi^2 = \sum_{i=1}^I \sum_{j=1}^J \frac{\left(n_{ij} - \frac{1}{n} n_{i.} n_{.j}\right)^2}{\frac{1}{n} n_{i.} n_{.j}} = n \cdot \sum_{i=1}^I \sum_{j=1}^J \frac{n_{ij}^2}{n_{i.} n_{.j}} - n \quad (\text{A.3})$$

$\chi^2$ -kvadrát rozdělení s  $(I-1)(J-1)$  stupni volnosti.

## LITERATURA

- ALLINGHAM, M. G. – SANDMO, A. (1972): Income Tax Evasion: A Theoretical Perspective. *Journal of Public Economics*, 1972, no. 1, pp. 323–338.
- BOADWAY, R. – MARCHAND, M. – PESTIEAU, P. (1994): Toward a Theory of the Direct-Indirect Tax Mix. *Journal of Public Economics*, vol. 55, 1994, pp. 71–88.
- COCHRAN, W. G. (1963): *Sampling techniques*. 2<sup>nd</sup> edition. New York, Wiley, 1963.
- CREMER, H. – GAHVARI, F. (1996): Tax Evasion and the Optimum General Income Tax. *Journal of Public Economics*, vol. 60, 1996, pp. 235–249.
- ENGEL, E. M. R. A. – HINES, J. R., Jr. (1999): Understanding Tax Evasion Dynamics. *National Bureau of Economic Research, Working Paper* no. 6903, January 1999.
- FEIGE, E. L. – MCGEE, R. T. (1983): Sweden's Laffer Curve: Taxation and the Unobserved Economy. *Scandinavian Journal of Economics*, vol. 85, 1983, pp. 499–519.
- FORTIN, B. – LACROIX, G. – MONTMARQUETTE, C. (2000): Are Underground Workers More Likely to be Underground Consumers? *Economic Journal*, October 2000, vol. 110 (466), pp. 838–60.
- HANOUSEK, J. – PALDA, F. (2002): Why People Evade Taxes in the Czech and Slovak Republics: A Tale of Twins. Prague, *CERGE-EI, Discussion Paper*, 2002, no. 85.
- HANOUSEK, J. – PALDA, F. (2003): Why People Evade Taxes in the Czech and Slovak Republic: A Tale of Twins. In: Belev, B. (ed.): *The Informal Economy in the EU Accession Countries: Size, Scope, Trends and Challenges in the Process of EU Enlargement*. Sofia, Center for the Study of Democracy, 2003, pp. 139–174.
- HANOUSEK, J. – PALDA, F. (2004): Quality of Government Services and the Civic Duty to Pay Taxes in the Czech and Slovak Republics and Other Transition Countries. *Kyklos*, vol. 57, 2004, no. 2, pp. 237–52.
- JUNG, Y. H. – SNOW, A. – TRANDEL, G. (1994): Tax Evasion and the Size of the Underground Economy. *Journal of Public Economics*, vol. 54, 1994, pp. 391–402.
- LUCAS, R. E. (1976): Econometric Policy Evaluation: A Critique. *Carnegie-Rochester Conference Series on Public Policy*, 1976, no. 1, pp. 19–46.
- McFADDEN, D. (1987): Regression-Based Specification Tests for the Multinomial Logit Model. *Journal of Econometrics*, vol. 34(1/2), 1987, pp. 63–82.
- REINGANUM, J. E. – WILDE, L. L. (1991): Equilibrium Enforcement and Compliance in the Presence of Tax Practitioners. *Journal of Law, Economics, and Organization*, 1991, no. 7, pp. 163–168.
- SCHNEIDER, F. – ENSTE, H. (2000): Informal Economies: Size, Causes, and Consequences. *Journal of Economic Literature*, vol. 38, 2000 (March), pp. 77–114.
- WATSON, H. (1985): Tax Evasion and Labor Markets. *Journal of Public Economics*, vol. 27, 1985, pp. 231–246.
- YANIV, G. (1994): Tax Evasion and the Income Tax Rate: A Theoretical Perspective. *Public Finance*, vol. 49, 1994, pp. 107–112.

## SUMMARY

JEL classification: H26, H43, K42, O17

Keywords: informal economy; shadow economy; tax evasion; transition

# **Estimating the Evolution of Tax Evasion: A Markov Chain Analysis Applied to the Czech Republic**

Jan HANOUSEK – Center for Research and Graduate Education-Economic Institute, Prague

(Jan.Hanousek@cerge-ei.cz)

Filip PALDA – Ecole Nationale d'Administration Publique, Montreal, Canada (Filip\_Palda@enap.quebec.ca)

The paper demonstrates how survey methods and retrospective questions can be put to effective use in the study of shadow-economy dynamics; specifically, the evolution of tax evasion in the Czech Republic. The authors measure the average individual's transition between the shadow and official economy, and use these transition probabilities to predict the evolution of tax evasion with the help of tax-evasion surveys taken in 2000, 2002, and 2004. The authors' results suggest that the shadow economy in the Czech Republic has flattened, and that a small decrease in the degree of evasion can be expected.